ELSEVIER

Discrete Applied Mathematics 119 (2002) 297–304

DISCRETE APPLIED MATHEMATICS

Note

Discrete convexity: convexity for functions defined on discrete spaces

Ümit Yüceer*

Department of Industrial Engineering, Eastern Mediterranean University, Gazimagusa, T.R.N.C., via Mersin 10 Turkey

Received 22 June 1998; received in revised form 17 October 2000; accepted 30 October 2000

Abstract

The concept of discrete convexity for a real-valued function defined on a discrete space is an extension of the convexity definition of continuous functions. The equivalence of discrete convexity to the conventional definition of increasing (non-decreasing) first forward differences of functions of single variables is established. A further extension of the discrete convexity with submodularity yields the concept of strong discrete convexity. A function with the property of strong discrete convexity has a positive semi-definite matrix of second forward differences.

Keywords: Discrete convexity; Strong discrete convexity; Matrix of second forward differences

1. Introduction

The concept of convexity is very important for continuous functions. It provides a good description of the behavior of a class of continuous functions. The convexity is identified by a positive-semi-definite Hessian matrix of second partial derivatives. There is also a strong relationship between convexity and sub/supermodularity. In addition, the local minimum of a strictly convex function is the global minimum.

There are several discrete functions arising from Operations Research and/or Management Science applications or as subproblems or dual problems in integer programming. A proper description of their behavior is required in optimizing such discrete problems. The classical definition states that a discrete function of a single variable is convex if its first forward differences are increasing or at least nondecreasing, as defined by Denardo [2], Fox [5] and many others in the literature. Favati and Tardella [3] investigate convexity properties of integer valued functions and produce some algorithmic

* Corresponding author.

E-mail address: umit.yuceer@emu.edu.tr (Ü. Yüceer).

0166-218X/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0166-218X(01)00191-3

results for nonlinear objective functions in integers. Murota [8] defines a concept of convexity for integer valued functions and investigates its relationship to submodularity. His work constitutes an important development towards conjugacy, subgradients, the Fenchel min-max duality and Lagrange duality for convex/nonconvex optimization. The discrete convexity of real–valued functions defined on a discrete space is proposed by Miller [6,7] as an extension of the definition of the convexity for continuous functions. One major contribution of this paper is to establish the equivalence of the classical definition of convexity of a separable discrete function to the discrete convexity for real-valued functions defined on a discrete space.

A marginal allocation algorithm for nonseparable functions, developed by Yüceer [9], generates undominated solutions under discrete convexity. Federgruen and Groenevelt [4] propose a concept of weak convexity (concavity) to implement a marginal allocation algorithm to solve some resource allocation problems. They fail to realize that their definition of weak convexity is actually a very strong condition.

A concept of strong discrete convexity is obtained by imposing additional conditions on a discretely convex function such as submodularity. A discretely convex function yields non-negative second forward differences in each component, and a symmetric matrix of second forward cross differences. A strongly convex function is proven to have a positive-semi-definite matrix of second forward differences. This has practical and computational implications. In particular, a separable function happens to be strongly discretely convex. Development of the concept of the strong discrete convexity and its implications is another contribution of this article.

Section 2 describes the concept of discrete convexity and establishes the main results. Section 3 presents the concept of strong discrete convexity and its implications. Section 4 furnishes an example and points out some further research areas.

2. Discrete convexity

Let S be a subspace of a discrete n-dimensional space. A function $f : S \to R$ is discretely convex if for all $x, y \in S$ and for $\alpha \in (0,1)$

$$\alpha f(x) + (1-\alpha) f(y) \geqslant \min_{u \in N(z)} f(u), \tag{1}$$

where $N(z) = \{u \in S : ||u-z|| < 1\}$, $z = \alpha x + (1-\alpha)y$, and $||u|| = \max_{1 \leqslant i \leqslant n}\{|u_i|\}$. This is a rather straightforward extension of the usual definition of convexity for continuous functions to discrete functions. It provides a description of the behavior of a class of discrete functions. The restriction of any continuous function to a discrete space does not necessarily yield a discretely convex function as illustrated in the following example which is an adaptation of Rosenbrook's function.

Let $f(y_1, y_2) = 25(2y_2 - y_1)^2 + \frac{1}{4}(2 - y_1)^2$ be defined from R^2 into the reals. This function is convex, since its Hessian is positive semi-definite at any point in R^2. The restriction of this function to the integers, however, is not discretely convex.

$x_2 \setminus x_1$	0	1	2
0	1	25.25	100
1	101	25.25	0

If $x=(0,0)$ and $y=(2,1)$ with $\alpha=0.5$, then $0.5f(0,0)+0.5f(2,1)=0.5$ but unfortunately $\min_{u\in N(1,0.5)} f(u) = 25.25$ where $N(1,0.5)=\{(1,0),(1,1)\}$.

A discrete rectangle is defined as $S=\{(x_1,x_2,\dots,x_n) : a_j \leqslant x_j \leqslant b_j \text{ for all } j=1,2,\dots,n\}$ where a_j and b_j are integers.

A function $f : D \to R$ is weighted if D is the convex hull of the discrete rectangle S and for $x \in D$, $f(x)=\sum_{u\in N(x)} w_u(x) f(u)$ where $w_u(x)$ are called the weights and satisfy the conditions $\sum_{u\in N(x)} w_u(x) = 1$ and $w_u(x) \geqslant 0$. The following lemma is borrowed from Miller [6,7].

Lemma 1. *If a weighted function $f : D \to R$ where $D \subset R^n$ is convex, then its restriction to discrete rectangle $S = D \cap I^n$ is discretely convex.*

The first forward difference of f in the direction of $e_i = i$th unit vector at a point $x \in S$ (provided that $x + e_i \in S$) is defined as follows.

$$\triangle_i f(x) = f(x+e_i) - f(x) \quad \text{for all } i = 1,\dots,n. \tag{2}$$

The second forward difference of f in the direction of e_i and e_j at a point $x \in S$ is given as follows.

$$\begin{aligned}\triangle_{ij} f(x) &= \triangle_i f(x+e_j) - \triangle_i f(x) \quad \text{for all } j = 1,\dots,n\\ &= f(x+e_i+e_j) - f(x+e_i) - f(x+e_j) + f(x) \quad \text{for all } i,j.\end{aligned}$$

Now, the equivalence of the discrete convexity to the commonly used definition of convexity for discrete functions which states that a discrete function is convex if its first forward differences are increasing (nondecreasing), see Denardo [2], Fox [5] or any others, will be established by the following theorem and corollary.

Theorem 1. *A discretely convex function of a single variable has its first forward differences increasing* (*nondecreasing*). *Conversely, if the first forward differences of a discrete function of a single variable are increasing* (*nondecreasing*), *then it is discretely convex.*

The proof of this theorem is straightforward. Setting $\alpha = 0.5$ in the definition immediately yields $\triangle f(x+1) \geq \triangle f(x)$. The converse requires showing the existence of a continuous weighted convex function as an extension of a discretely convex function.

Corollary 1. *A separable function is discretely convex if and only if it is discretely convex in each component.*

Proof. Let $f(x) = \sum_{j=1}^{n} f_j(x_j)$ be a separable discretely convex function. Then the first forward differences $\triangle_j f(x) = \triangle f_j(x_j)$ are increasing for each $j = 1, 2, \ldots, n$, hence each $f_j(x_j)$ is discretely convex. Conversely, if each $f_j(x_j)$ is discretely convex, then for $\alpha \in (0, 1)$ the relationship $\alpha f_j(x_j) + (1 - \alpha) f_j(y_j) \geqslant \min_{u_j \in N(z_j)} f_j(u_j)$ holds for all $j = 1, 2, \ldots$. Summing yields the following relationship:

$$\alpha \sum_{j=1}^{n} f_j(x_j) + (1 - \alpha) \sum_{j=1}^{n} f_j(y_j) \geqslant \sum \min_{u_j \in N(z_j)} f_j(u_j). \tag{3}$$

The right-hand side of (3) becomes equal to $\min_{u \in N(z)} \sum_{j=1}^{n} f_j(u_j)$ by defining $z = (z_1, z_2, \ldots, z_n)$ and $u = (u_j, u_2, \ldots, u_n)$. Finally, $\alpha f(x) + (1 - \alpha) f(y) \geqslant \min_{u \in N(z)} f(u)$. This establishes the equivalence of discrete convexity to the commonly used definition for discrete separable functions defined as increasing (nondecreasing) first forward differences in the literature. □

3. Strong discrete convexity

Let $f : S \to R$ be a discretely convex function. The function f has strong discrete convexity if it satisfies the following conditions:

1. $f(x + u) + f(x) \geqslant f(x \vee u) + f(x \wedge u)$ where $u = (u_1, \ldots, u_n) \neq 0$, $u_i = 0, -1, +1$ for each $i = 1, 2, \ldots, n$ and $x \vee u = (\max\{x_i, x_i + u_i\})$, $x \wedge u = (\min\{x_i, x_i + u_i\})$. This is a local submodularity condition.
2. $\sum_{j=1}^{n} \triangle_j f(x) \leqslant \sum_{j=1}^{n} \triangle_j f(x + e_i)$ for all $i = 1, 2, \ldots, n$, or $\sum_{j=1}^{n} \triangle_{ij} f(x) \geq 0$.

Corollary 2. *A separable discretely convex function has the property of strong discrete convexity.*

The proof of this corollary is rather trivial. Therefore, a separable discretely convex function is always strongly discretely convex. This result implies that separability itself is a stronger condition.

A discretely convex function has $\triangle_{ii} f(x) \geqslant 0$ for all $i = 1, 2, \ldots, n$. This is an equivalent statement to increasing first forward differences. The matrix of second forward differences is symmetric with nonnegative entries on the diagonal.

Theorem 2. *The matrix of second forward differences of a strongly discretely convex function is positive semi-definite.*

Proof. A strongly discretely convex function has $\triangle_{ij} f(x) \leqslant 0$ for all $j \neq i$ as a result of local submodularity. Consequently, $f(x + e_i) + f(x + e_j) \geqslant f(x + e_i + e_j) + f(x)$ holds and the following relationship is obtained for all $j \neq i$.

$$\triangle_{ij} f(x) = f(x + e_i + e_j) - f(x + e_i) - f(x + e_j) + f(x) \leqslant 0. \tag{4}$$

The matrix of second forward differences is diagonally dominant if the following statement holds for all $i = 1, 2, \ldots, n$:

$$\triangle_{ii} f(x) \geqslant \sum_{j \neq i} |\triangle_{ij} f(x)|. \tag{5}$$

Since $\triangle_{ij} f(x) \leqslant 0$, then the term

$$T = \triangle_{ij} f(x) - \sum_{j \neq i} |\triangle_{ij} f(x)| \tag{6}$$

can be expressed as follows, and after some algebraic manipulations, the result is obtained.

$$T = \triangle_{ij} f(x) + \sum_{j \neq i} \triangle_{ij} f(x), \tag{7}$$

$$T = \sum_{j=1}^{n} \triangle_j f(x + e_j) - \sum_{j=1}^{n} \triangle_j f(x). \tag{8}$$

By condition (2), this term is nonnegative; hence, the matrix of second forward differences is diagonally dominant. If a symmetric real matrix (Hermitian) with nonnegative diagonal entries is diagonally dominant, then it is positive semi-definite, [1]. $\square$

Only a partial converse of this theorem holds. If the matrix of second forward differences of a discretely convex function together with the local submodularity condition is positive semi–definite, then the function has the property of strong discrete convexity. The discrete convexity property together with the positive semi–definiteness of the matrix of second forward differences implies that $\triangle_{ii} f(x) \geq 0$ for all $i = 1, 2, \ldots, n$ and $\sum_{j=1} \triangle_{ij} f(x) \geq 0$. Hence, the second condition of the strong discrete convexity is also satisfied. This establishes an analogous property (even though highly restrictive) for the functions with the property of strong discrete convexity to the continuous convex functions. Unfortunately, no inferences about the discrete convexity of a function can be derived from the matrix of second forward differences of discrete functions. The former result has some practical and computational implications. It provides a simple check on whether a discretely convex function also has the property of strong discrete convexity at some point. An example of a discretely convex function will be furnished in the following section to illustrate these concepts.

4. Conclusions and an example

An example of a discretely convex function is furnished below to illustrate the concepts discussed and developed in this research:

$$f(x) = \sum_{k=0}^{\infty} \left(1 - \prod_{j=1}^{n} \beta_j (x_j + k) \right) + \lambda \sum_{j=1}^{n} c_j x_j, \tag{9}$$

where $\lambda > 0$, $\beta_j(.)$ is the cumulative Poisson sums with a mean of μ_j, $c_j > 0$ and x_j is a nonnegative integer for all $j = 1, 2, \ldots, n$. The infinite sum is convergent, since the summand is negligible for all sufficiently large k. The discrete convexity of (9) is proved by Miller [6] or [7]. The strong discrete convexity of (9) will be shown here.

The first and second forward differences of $f(x)$ will be calculated as follows.

$$\triangle_i f(x) = \lambda c_i - \sum_{k=0}^{\infty} \phi_i(x_i + k + 1) \prod_{t \neq i} \beta_t(x_t + k), \tag{10}$$

$$\triangle_{ii} f(x) = \sum_{k=0}^{\infty} (\phi_i(x_i + k + 1) - \phi_i(x_i + k + 2)) \prod_{t \neq i} \beta_t(x_t + k), \tag{11}$$

$$\triangle_{ij} f(x) = -\sum_{k=0}^{\infty} \phi_i(x_i + k + 1)\phi_j(x_j + k + 1) \prod_{t \neq i,j} \beta_t(x_t + k) \text{ for } i \neq j, \tag{12}$$

where $\phi(.)$ is the individual Poisson terms. Clearly, $\triangle_{ij} f(x) \leq 0$ for $i \neq j$ and $\triangle_{ii} f(x) \geq 0$ by the discrete convexity property for all $i = 1, 2, \ldots, n$.

Theorem 3. *The function $f(x)$ of (9) has the property of strong discrete convexity.*

It is very easy to show that the function $f(x)$ is locally submodular and the rest of the proof is given in the appendix. As an illustration, we consider an example containing n=5 components with (μ_j)=$(2.1, 1.5, 1.2, 5.0, 3.5)$ and (c_j)=$(2980, 1751, 462, 1500, 345)$. The matrix of second forward differences at the point $(3, 2, 3, 6, 6)$ is given below. It is symmetric with all the diagonal entries which are positive, and it is diagonally dominant. Hence it is positive semi-definite.

$$H(3,2,3,6,6) = \begin{pmatrix} +0.03465 & -0.00164 & -0.00021 & -0.00255 & -0.00061 \\ -0.00164 & +0.03882 & -0.00023 & -0.00277 & -0.00067 \\ -0.00021 & -0.00023 & +0.00482 & -0.00034 & -0.00009 \\ -0.00255 & -0.00277 & -0.00034 & +0.05976 & -0.00106 \\ -0.00061 & -0.00067 & -0.00009 & -0.00106 & +0.01380 \end{pmatrix}.$$

In this article, a convexity concept for a real–valued function defined on a discrete space and its extensions are discussed. There are some resource allocation problems in real life with such objective functions or constraints. Some concept of convexity is essential for developing algorithms to solve such problems. It remains to investigate further the relationship between the discrete convexity and submodularity. It may also be very interesting to investigate duality relationships, conjugacy, and the Fenchel min–max duality for discretely convex functions defined on discrete spaces.

Acknowledgements

The author wishes to thank an anonymous referee for his helpful comments and suggestions.

Appendix

The function $f(x)$ of (9) satisfies the second condition of the strong discrete convexity. This statement is expressed as a theorem below and its proof is given.

Theorem A.1. *The function $f(x)$ of* (9) *satisfies the condition* $\sum_{j=1}^{n} \triangle_j f(x) \leqslant \sum_{j=1}^{n} \triangle_j f(x+e_i)$ *or rather* $\sum_{j=1}^{n} \triangle_{ij} f(x) \geqslant 0$ for all $i = 1, 2, \ldots, n$.

Proof. Using expressions (11) and (12) the term $T = T_1 - T_2 - T_3$ where T_1, T_2, and T_3 are given below, is obtained.

$$T_1 = \sum_{k=0}^{\infty} \phi_i(x_i + k + 1) \prod_{t \neq i} \beta_t(x_t + k), \tag{A.1}$$

$$T_2 = \sum_{k=0}^{\infty} \phi_i(x_i + k + 2) \prod_{t \neq i} \beta_t(x_t + k), \tag{A.2}$$

$$T_3 = \sum_{j \neq i} \sum_{k=0}^{\infty} \phi_i(x_i + k + 1)\phi_j(x_j + k + 1) \prod_{t \neq i,j} \beta_t(x_t + k). \tag{A.3}$$

The infinite sum in T_1, T_2, and T_3 is convergent. The term $T_1 - T_2 = \triangle_{ii} f(x)$ is nonnegative by discrete convexity. Rearranging the index of the sum in T_2 yields the following:

$$T_2 = \sum_{k=1}^{\infty} \phi_i(x_i + k + 1) \prod_{t \neq i} \beta_t(x_t + k - 1). \tag{A.4}$$

Consequently, the term $T_{12} = T_1 - T_2$ can be expressed as follows:

$$T_{12} = \sum_{k=1}^{\infty} \phi_i(x_i + k + 1) \left(\prod_{t \neq i} \beta_t(x_t + k) - \prod_{t \neq i} \beta_t(x_t + k - 1) \right)$$

$$+ \phi_i(x_i + 1) \prod_{t \neq i} \beta_t(x_t). \tag{A.5}$$

Clearly, if the term $\prod_{t \neq i} \beta_t(x_t + k) - \prod_{t \neq i} \beta t(x_t + k - 1)$ is nonnegative, then each term in the infinite sum is nonnegative. Since $\beta_t(x_t + k) = \beta_t(x_t + k - 1) + \phi_t(x_t + k)$ for each $k \geqslant 1$ the following relationship is obtained:

$$\prod_{t \neq i} \beta_t(x_t + k) = \prod_{t \neq i} (\beta_t(x_t + k - 1) + \phi_t(x_t + k)). \tag{A.6}$$

Expansion of the term on the right-hand side and an algebraic manipulation yield the following ($O(\xi)$ which represents sum of relatively small but nonnegative terms):

$$\prod_{t \neq i} \beta_t(x_t + k) = \prod_{t \neq i} \beta_t(x_t + k - 1) + \sum_{j \neq i} \phi_j(x_j + k) \prod_{t \neq i,j} \beta_t(x_t + k - 1) + O(\phi_j), \tag{A.7}$$

$$1-\frac{\prod_{t\neq i}\beta_t(x_t+k-1)}{\prod_{t\neq i}(\beta_t(x_t+k))}=\sum_{j\neq i}\frac{\phi_j(x_j+k)}{\beta_j(x_j+k)}+O(\phi_j). \tag{A.8}$$

On the other hand, T_3 can be expressed after an algebraic manipulation as follows:

$$T_3=\sum_{k=0}^{\infty}\phi_i(x_i+k+1)\prod_{t\neq i}\beta_t(x_t+k)\sum_{j\neq i}\frac{\phi_j(x_j+k+1)}{\beta_j(x_j+k)}. \tag{A.9}$$

If $\mu_j \leqslant x_j+1$ for all $j=1,2,\ldots,n$ then $\phi_j(x_j+k+1)\leqslant\phi_j(x_j+k)$ which implies the relationship for all $k\geqslant 1$ given below.

$$\sum_{j\neq i}\frac{\phi_j(x_j+k)}{\beta_j(x_j+k)}\geqslant\sum_{j\neq i}\frac{\phi_j(x_j+k+1)}{\beta_j(x_j+k)}. \tag{A.10}$$

This relationship implies that each summand in T_{12} is greater than or equal to each summand in T_3 for all $k\geqslant 1$. If $k=0$, then $\sum_{j\neq i}\phi_j(x_j+1)/\beta_j(x_j)\leqslant 1$ holds for sufficiently large x_j for all $j=1,2,\ldots,n$ which implies that $T_{12}\geqslant T_3$ or $T\geqslant 0$. Thus, condition (2) is satisfied for a sufficiently large $(x_1,x_2,\ldots,x_n)$. This completes the proof and the matrix of second forward differences is positive semi-definite. □

References

[1] R. Bellman, Introduction to Matrix Analysis, 2nd Edn., McGraw-Hill, New York, 1970.

[2] E.V. Denardo, Dynamic Programming, Prentice-Hall, Englewood cliffs, NJ, 1982.

[3] P. Favati, F. Tardella, Convexity in nonlinear integer programming, Ricerca Operativa 53 (1990) 3–44.

[4] A. Federgruen, H. Groenevelt, The greedy procedure for resource allocation problems: necessary and sufficient conditions for optimality Oper. Res. (34) 1986 909–918.

[5] B. Fox, Discrete optimization via marginal analysis, Management Sci. 13 (1966) 210–216.

[6] B.L. Miller, Unconstrained Optimization in Integers RM-6165-PR, The RAND Corporation, Santa Monica, California, 1970.

[7] B.L. Miller, On minimizing nonseparable functions defined on the integers with an inventory application, SIAM J. Appl. Math. 21 (1971) 1–15.

[8] K. Murota, Discrete convex analysis, Math. Programming 83 (1998) 313–371.

[9] U. Yüceer, Marginal allocation algorithm for nonseparable functions, INFOR 37 (1999) 97–113.